

Actividad 6 (Regresión No Lineal)



Analítica de datos y herramientas de inteligencia artificial II (Gpo 501)

IIS

PROFESOR:

Alfredo García Suárez

PRESENTA:

Ricardo Saldaña Muñoz | A01707504

18 de octubre de 2023
Campus Querétaro

Introducción:

En el presente documento se presenta un reporte comparativo de los análisis generados por cada una de las variables objetivo. Entre el análisis podremos observar de cada variable objetivo, con cuales variables tiene un alto coeficiente de determinación y un alto coeficiente de correlación y qué es lo que quiere decir que tengan ese valor.

Las variables objetivas analizadas fueron:

- Riesgo
- Score_buro
- Porc_eng
- Limite_credito

Análisis de la variable objetivo – Riesgo

A continuación se presenta una tabla en la cual se muestran las variables que se analizaron entre si mediante el método de Regresión No Lineal, se muestra la función utilizada por modelo, y además se muestra tanto la R2 como el coeficiente de correlación:

	Variable Dependiente	Variable Independiente	Funcion Utilizada	R2	Correlacion
0	riesgo	porc_eng	$(a * x^{**2} + b) / c * x$ (Función cuadrática)	0.206305	0.454209
1	riesgo	limite_credito	$a/b * x^{**2} + c * x$ (Función polinomial inversa)	0.331664	0.575903
2	riesgo	puntos	$1/a * x^{**2}$ (Función cuadrática inversa)	-0.090863	0.301434
3	riesgo	enganche	$a * x + b * x + c * x$ (Función lineal con pro...	0.063954	0.252891
4	riesgo	pagos_realizados	$a * np.abs(x) + b * x + c$ (Función Valor absol...	0.057966	0.240760

Conclusión:

Riesgo vs Porc_eng:

El R2 de 0.206305 significa que aproximadamente el 20.63% de la variabilidad en la variable dependiente "Riesgo" puede ser explicada por el porcentaje de enganche ("porc_eng"). La correlación (r) de 0.454209 indica una relación moderada positiva entre estas dos variables. Esto sugiere que a medida que el porcentaje de enganche aumenta, el riesgo tiende a aumentar, pero la relación no es muy fuerte.

Riesgo vs limite_credito:

El R2 de 0.331664 significa que aproximadamente el 33.17% de la variabilidad en la variable dependiente "Riesgo" puede ser explicada por el límite de crédito ("limite_credito"). La correlación (r) de 0.575903 indica una correlación positiva moderada a fuerte entre estas dos variables. Esto sugiere que a medida que el límite de crédito aumenta, el riesgo tiende a aumentar, y la relación es más fuerte que en el caso anterior.

Riesgo vs Puntos:

En este caso, el R2 es negativo, lo que implica que el modelo de regresión no es adecuado para explicar la variabilidad en la variable dependiente "Riesgo" a partir de la variable independiente "puntos". Además, la correlación (r) de 0.301434 indica una correlación positiva, pero no es muy fuerte. La relación es débil y negativa debido al R² negativo, lo que sugiere que los puntos no son un buen predictor del riesgo.

Riesgo vs Enganche:

El R2 de 0.063954 significa que aproximadamente el 6.40% de la variabilidad en la variable dependiente "Riesgo" puede ser explicada por el enganche. La correlación (r) de 0.252891 indica una relación positiva, pero es relativamente débil. Esto sugiere que hay una relación positiva entre el enganche y el riesgo, pero el enganche no es un predictor muy fuerte del riesgo.

Riesgo vs Pagos_realizados:

El R2 de 0.057966 significa que aproximadamente el 5.80% de la variabilidad en la variable dependiente "Riesgo" puede ser explicada por los pagos realizados. La correlación (r) de 0.24076 indica una relación positiva, pero es relativamente débil. Esto sugiere que hay una relación positiva entre los pagos realizados y el riesgo, pero los pagos realizados no son un predictor muy fuerte del riesgo.

Análisis de la variable objetivo – Score_buro

A continuación se presenta una tabla en la cual se muestran las variables que se analizaron entre sí mediante el método de Regresión No Lineal, se muestra la función utilizada por modelo, y además se muestra tanto la R2 como el coeficiente de correlación:

	Variable Dependiente	Variable Independiente	Funcion Utilizada	R2	Correlacion
0	score_buro	status	$(a * x^{**2} + b) / c * x$ (Función cuadrática)	0.050314	0.224308
1	score_buro	pagos_realizados	$a/b * x^{**2} + c * x$ (Función polinomial inversa)	-0.340364	0.583407
2	score_buro	puntos	$1/a * x^{**2}$ (Función cuadrática inversa)	-0.992311	0.996148
3	score_buro	semana_actual	$a * x + b * x + c * x$ (Función lineal con pro...)	-0.698949	0.836032
4	score_buro	limite_credito	$a * np.abs(x) + b * x + c$ (Función Valor absol...)	0.023516	0.153349

Conclusión:

Score_buro vs Status:

El R2 de 0.050314 significa que aproximadamente el 5.03% de la variabilidad en el "Score de Buro" puede ser explicada por la variable "Status". La correlación (r) de 0.224308 indica una correlación positiva débil entre estas dos variables.

Score_buro vs Pagos_realizados:

El R2 negativo (-0.340364) sugiere que el modelo de regresión no es adecuado para explicar la variabilidad en el "Score de Buro" a partir de la variable "Pagos Realizados". Sin embargo, la correlación (r) de 0.583407 indica una correlación negativa moderada entre estas dos variables. Esto significa que a medida que los pagos realizados aumentan, el "Score de Buro" tiende a disminuir.

Score_buro vs Puntos:

El R2 muy negativo (-0.992311) sugiere que el modelo de regresión no es adecuado para explicar la variabilidad en el "Score de Buro" a partir de la variable "Puntos". Sin embargo, la correlación (r) de 0.996148 indica una correlación extremadamente alta y positiva entre estas dos variables. Esto sugiere que los puntos y el "Score de Buro" están altamente relacionados, pero el modelo de regresión no es apropiado para modelar esta relación.

Score_buro vs Semana_actual:

El R2 negativo (-0.698949) indica que el modelo de regresión no es adecuado para explicar la variabilidad en el "Score de Buro" a partir de la variable "Semana Actual". Sin embargo, la correlación (r) de 0.836032 indica una correlación positiva muy fuerte entre estas dos variables. Esto sugiere que el "Score de Buro" y la "Semana Actual" están altamente relacionados, pero el modelo de regresión no es adecuado para modelar esta relación.

Score_buro vs Limite_credito:

El R2 de 0.023516 indica que aproximadamente el 2.35% de la variabilidad en el "Score de Buro" puede ser explicada por el "Límite_Crédito". La correlación (r) de 0.153349 indica una correlación positiva débil entre estas dos variables. Esto sugiere que el "Límite de Crédito" y el "Score de Buro" están débilmente relacionados en este caso.

Análisis de la variable objetivo – Porc_eng

A continuación, se presenta una tabla en la cual se muestran las variables que se analizaron entre si mediante el método de Regresión No Lineal, se muestra la función utilizada por modelo, y además se muestra tanto la R2 como el coeficiente de correlación:

	Variable Dependiente	Variable Independiente	Funcion Utilizada	R2	Correlacion
0	porc_eng	enganche	$(a * x^{**2} + b) / c * x$ (Función cuadrática)	0.076085	0.275836
1	porc_eng	costo_total	$a/b * x^{**2} + c * x$ (Función polinomial inversa)	-0.485726	0.696941
2	porc_eng	puntos	$a * np.abs(x) + b * x + c$ (Función Valor absol...	0.186863	0.432276
3	porc_eng	riesgo	$a * np.abs(x) + b * x + c$ (Función Valor abso...	0.297492	0.545428
4	porc_eng	limite_credito	$a * np.abs(x) + b * x + c$ (Función Valor absol...	0.442110	0.664914

Conclusión:

Porc_eng vs Enganche:

El R2 de 0.076085 significa que aproximadamente el 7.61% de la variabilidad en el "Porcentaje de Enganche" puede ser explicada por el "Enganche". La correlación (r) de 0.275836 indica una correlación positiva débil entre estas dos variables. Esto sugiere que a medida que el monto del enganche aumenta, el porcentaje de enganche tiende a aumentar, pero la relación no es muy fuerte.

Porc_eng vs Costo_total:

El R2 negativo (-0.485726) sugiere que el modelo de regresión no es adecuado para explicar la variabilidad en el "Porcentaje de Enganche" a partir del "Costo Total". Sin embargo, la correlación (r) de 0.696941 indica una correlación negativa fuerte entre estas dos variables. Esto significa que a medida que el costo total aumenta, el porcentaje de enganche tiende a disminuir.

Porc_eng vs Puntos:

El R2 de 0.186863 indica que aproximadamente el 18.69% de la variabilidad en el "Porcentaje de Enganche" puede ser explicada por los "Puntos". La correlación (r) de 0.432276 indica una correlación positiva moderada entre estas dos variables. Esto sugiere que a medida que los puntos aumentan, el porcentaje de enganche tiende a aumentar, y la relación es moderada.

Porc_eng vs Riesgo:

El R2 de 0.297492 significa que aproximadamente el 29.75% de la variabilidad en el "Porcentaje de Enganche" puede ser explicada por el "Riesgo". La correlación (r) de 0.545428 indica una correlación positiva moderada entre estas dos variables. Esto sugiere que a medida que el riesgo aumenta, el porcentaje de enganche tiende a aumentar, y la relación es moderada.

Porc_eng vs Limite_credito:

El R2 de 0.44211 significa que aproximadamente el 44.21% de la variabilidad en el "Porcentaje de Enganche" puede ser explicada por el "Límite de Crédito". La correlación (r) de 0.664914 indica una correlación positiva fuerte entre estas dos variables. Esto sugiere que a medida que el límite de crédito aumenta, el porcentaje de enganche tiende a aumentar, y la relación es fuerte.

Análisis de la variable objetivo – Limite_credito

A continuación se presenta una tabla en la cual se muestran las variables que se analizaron entre sí mediante el método de Regresión No Lineal, se muestra la función utilizada por modelo, y además se muestra tanto la R2 como el coeficiente de correlación:

	Variable Dependiente	Variable Independiente	Funcion Utilizada	R2	Correlacion
0	limite_credito	porc_eng	$(a * x^2 + b) / c * x$ (Función cuadrática)	-0.377990	0.614809
1	limite_credito	riesgo	$a * \text{np.abs}(x) + b * x + c$ (Función Valor absol...	0.543357	0.737127
2	limite_credito	puntos	$a * \text{np.abs}(x) + b * x + c$ (Función Valor absol...	0.268698	0.518361
3	limite_credito	semana_actual	$a * \text{np.abs}(x) + b * x + c$ (Función Valor absol...	0.115471	0.339811
4	limite_credito	pagos_realizados	$a * \text{np.abs}(x) + b * x + c$ (Función Valor absol...	0.061648	0.248289

Conclusión:

Limite_credito vs Porc_eng:

El R2 negativo (-0.37799) sugiere que el modelo de regresión no es adecuado para explicar la variabilidad en el "Límite de Crédito" a partir del "Porcentaje de Enganche". Sin embargo, la correlación (r) de 0.614809 indica una correlación positiva moderada entre estas dos variables. Esto sugiere que a

medida que el porcentaje de enganche aumenta, el límite de crédito tiende a aumentar, pero el modelo de regresión no captura bien esta relación.

Limite_credito vs Riesgo:

El R2 de 0.543357 significa que aproximadamente el 54.34% de la variabilidad en el "Límite de Crédito" puede ser explicada por el "Riesgo". La correlación (r) de 0.737127 indica una correlación positiva fuerte entre estas dos variables. Esto sugiere que a medida que el riesgo aumenta, el límite de crédito tiende a aumentar, y la relación es fuerte.

Limite_credito vs Puntos:

El R2 de 0.268698 significa que aproximadamente el 26.87% de la variabilidad en el "Límite de Crédito" puede ser explicada por los "Puntos". La correlación (r) de 0.518361 indica una correlación positiva moderada entre estas dos variables. Esto sugiere que a medida que los puntos aumentan, el límite de crédito tiende a aumentar, y la relación es moderada.

Limite_credito vs Semana_actual:

El R2 de 0.115471 significa que aproximadamente el 11.55% de la variabilidad en el "Límite de Crédito" puede ser explicada por la "Semana Actual". La correlación (r) de 0.339811 indica una correlación positiva, pero es relativamente débil. Esto sugiere que la "Semana Actual" y el "Límite de Crédito" están débilmente relacionados.

Limite_credito vs Pagos_realizados:

El R2 de 0.061648 significa que aproximadamente el 6.16% de la variabilidad en el "Límite de Crédito" puede ser explicada por los "Pagos Realizados". La correlación (r) de 0.248289 indica una correlación positiva, pero es relativamente débil. Esto sugiere que los "Pagos Realizados" y el "Límite de Crédito" están débilmente relacionados.